S1 Material. Equations used for batch effect removal.

We assume that measured and log2-transformed expression values of gene *i* in sample *j* of batch X can be expressed in a general form as follows:

$$x_{i,j} = x'_{i,j} + b_{i,j}^X + \varepsilon_{i,j}^X$$

$x'_{i,j}$ represents the actual gene expression, $b_{i,j}^X$ the batch effect term and $\varepsilon_{i,j}^X$ reflects noise. We estimated $b_{i,j}^X$ by fitting linear regression models. Therefore, samples were assigned to either batch 1 or batch 2 representing baseline and follow-up measurements respectively. Subsequently, $b_{i,j}^X$ was subtracted from the observed gene expression $x_{i,j}$. For the linear mixed model based batch effect removal, a random variable was used to model sample pairs across batches.